

IOWA STATE UNIVERSITY

Digital Repository

Economics Presentations, Posters and Proceedings

Economics

10-2014

Is Government Involvement Really Necessary: Implications for Systemic Risk and Crop Reinsurance Contracts

Xiaoguang Feng

Iowa State University, xgfeng@iastate.edu

Dermot J. Hayes

Iowa State University, dhayes@iastate.edu

Follow this and additional works at: http://lib.dr.iastate.edu/econ_las_conf



Part of the [Agribusiness Commons](#), [Agricultural and Resource Economics Commons](#), [Agricultural Economics Commons](#), [Economics Commons](#), and the [Insurance Commons](#)

Recommended Citation

Feng, Xiaoguang and Hayes, Dermot J., "Is Government Involvement Really Necessary: Implications for Systemic Risk and Crop Reinsurance Contracts" (2014). *Economics Presentations, Posters and Proceedings*. Paper 4.
http://lib.dr.iastate.edu/econ_las_conf/4

This Conference Proceeding is brought to you for free and open access by the Economics at Digital Repository @ Iowa State University. It has been accepted for inclusion in Economics Presentations, Posters and Proceedings by an authorized administrator of Digital Repository @ Iowa State University. For more information, please contact digirep@iastate.edu.

**Is Government Involvement Really Necessary:
Implications for Systemic Risk and Crop
Reinsurance Contracts**

Xiaoguang Feng
Graduate Assistant
Department of Economics
Iowa State University
xgfeng@iastate.edu

Dermot J. Hayes
Pioneer Chair of Agribusiness
Professor of Economics
Professor of Finance
Iowa State University
dhayes@iastate.edu

***Selected Paper prepared for presentation at the Agricultural & Applied Economics Association's
Crop Insurance and the 2014 Farm Bill Symposium, Louisville, KY, October 8-9, 2014.***

Copyright 2014 by Xiaoguang Feng, Dermot J. Hayes. All rights reserved. Readers may make verbatim copies of this document for non-commercial purposes by any means, provided that this copyright notice appears on all such copies.

1 Introduction

Agriculture is subject to substantial systemic risk of crop yield losses due to widespread natural disasters. The systemic risk has been a major obstacle for the development of private crop insurance markets. Driven by spatially correlated weather events, crop losses are highly correlated within a certain area. As a result, the portfolio insurance risk associated with the crop losses has been raised far above what it would be if individual losses were independent, as proposed by Miranda and Glauber (1997). For example, Miranda and Glauber (1997) find that the portfolio risk faced by U.S. crop insurers is about ten times larger than that of conventional insurance lines. Large portfolio risk requires high premium rates to cover the cost of bearing the systemic portfolio risk unless the cost is subsidized. Some national governments, such as the U.S., are willing to provide subsidies and reinsurance for crop insurance policies so that they are affordable to farmers. In this way, the cost of bearing the systemic risk has been transferred to governments. For those countries where there are no government subsidies, private crop insurers would have to charge high premiums, in order to hold large enough reserves for the potential systemic loss or purchase expensive international reinsurance. In this way, the cost of bearing the systemic risk is actually passed onto farmers eventually. Consequently, farmers are either buying extremely expensive insurance to get insured, or being exposed to huge crop loss risks.

Systemic risk in agriculture has made crop insurance markets not so effective or independent as other insurance lines, imposing negative impacts on the welfare of farmers as well as the development of farm sector. To be able to provide crop insurance, private insurers need instruments to transfer the systemic portfolio insurance risk, either in terms of government subsidies or high premium rates. Without sufficient subsidies, crop insurance is losing its effect as farmers would hardly be willing to pay the extremely high premiums. Then in

the face of a shock that causes large yield losses, farmers would have to resort to survival strategies that could undermine their resilience and food security in the long-run. As a result, farmers are trapped in chronic destitution and agriculture remains underdeveloped. Even if government subsidization is available, farm sector is still affected by the systemic risk because the funds used as subsidies could be allocated to other development activities if there were no systemic risk. Therefore, if the systemic risk could be managed with more efficient tools, farm sector would be substantially benefited from it with improved farmers' welfare and enhanced agricultural development.

This document proposes a potential solution for the systemic risk problem in agriculture. To better deal with the systemic risk, a central question is that whether the systemic risk in agriculture is inherently non-diversifiable as its name suggests, or it is just because the risk pool is too small to be diversified, that is, the so-called systemic risk can in fact become diversifiable if the risk pool is large enough. Some studies have investigated the effectiveness of diversifying the systemic risk by enlarging the risk pool. Wang and Zhang (2013) suggest that the correlation of crop yields is observed to be decreasing with distance, so risk pooling is effective and a private crop insurance market is possible in the U.S.. However, this study ignores an important feature of systemic risk in agriculture — the risk is state-dependent. Specifically, the correlation of crop yields tends to be much stronger during extreme weather than in normal years (Goodwin, 2001). Therefore, the assumption of linear correlation of crop yields does not seem appropriate because it may understate the magnitude of systemic risk in extreme years. To measure the state-dependent systemic risk more correctly, Xu et al. (2010) and Okhrin, Odening, and Xu (2012) apply copulas to model the correlation of crop yields. However, these two studies have found that the risk pool is not large enough to support a viable private index-based insurance within Germany and China, respectively.

To consider a large enough risk pool, we propose an investigation of the diversification effect across multiple crops and multiple countries. If there was evidence that systemic risk could be diversified within the pool, then creating such a risk pool would bring many benefits to farm sector and agriculture. First, the risk pooling would facilitate the optimal use of crop insurance. As the risks in the pool were no longer being systemic, insurers would be able to effectively pool the risk and reduce the insurance costs of all the individual participants. Then the funds used to subsidize the insurance costs could be redirected to other agricultural development activities. In addition, as the systemic risk could be diversified in the pool, insurers would no longer need to hold large reserves by charging extremely high premiums. Then with low premium rates farmers would be able and willing to purchase crop insurance that helps them avoid being exposed to large yield losses risk. With the risk pool established across the world, it would help reduce vulnerability, improve resilience, and increase welfare for all the farmers, as well as protect food security and human, social, and economic development, especially for the least developed countries in the world.

The effectiveness of the risk pool described above depends on two factors. First, risk pooling across crops and countries has the ability to remove the systemic nature of risks in the pool, that is, risks are diversifiable in the risk pool. Second, there are many enough participants so that the systemic risk in the pool can be well-diversified because of law of large numbers. In this proposal, we attempt to perform a preliminary study on two large agricultural producing countries the U.S. and China and five major crops produced in these two countries. We consider a synthetic area-yield insurance portfolio across both countries and all crops at state/province-level as it was shown that the risk-reducing effect is the most significant if area-yield insurance is provided at state-level (Miranda and Glauber, 1993). If the study finds that the systemic nature of the insurance portfolio risk can be removed by

the risk pooling across the crops and countries, the proposed study can be performed on more crops and regions to see if systemic risk in agriculture can be finally eliminated.

To quantify the extent of area-yield insurance portfolio risk, it is essential to measure the correlation among the area-yields. In this study, a copula-based approach is proposed to model the joint behavior of the yield variables. As linear correlation cannot fully represent the state-dependent correlation structure for crop yields, copulas can be a nice alternative as they allow for greater flexibility in modeling correlations, such as tail dependence. By copula modeling, more accurate measurements are obtained for the correlation of yield variables as well as the risk in an insurance portfolio. Many copula models have been applied for multivariate modeling in agriculture, including some advanced copulas such as vine copulas and hierarchical Archimedean copulas (HAC) (Goodwin, 2012; Xu et al., 2010). Given the potentially high dimensions, our study applies the hierarchical Kendall copula (HKC), a recent innovation in copulas, to estimate the correlation among yield variables. Compared to vine copulas and HAC, HKC allows for both flexibility and parsimony in modeling the joint distribution of highly dimensional variables. The hierarchical structure of the HKC ensures that it is parsimonious in terms of the numbers of copula parameters. Meanwhile, the choice of the basic copula at each hierarchical level is not limited to any copula class, which means that HKC is flexible in modeling various kinds of correlation structures, such as asymmetric tail dependences.

The estimation of the HAC involves a sequential estimating procedure. Parameters of the basic copulas at the lowest hierarchical level are estimated first. Parameters of the copulas at higher levels are estimated consecutively by plugging-in the estimates of the copula parameters from lower levels. To increase efficiency and accuracy of the estimation, we use a Bayesian approach to take into account estimation risk for the sequential estimation

of the HAC parameters. This approach avoids accumulating large estimation errors at each sequential estimating step, which is especially important given the high dimensions of the HKC and a rather short time series of available yield observations. With the estimation results, the systemic risk associated with an insurance portfolio can be assessed from the predictive distribution of joint insurance losses.

The rest of the paper is organized as follows. We start with a review of previous work on the relevance of systemic risk in crop yields and copula modeling of correlation structures. The analytical model is then discussed in Section 3. Section 4 presents an application of the model to empirical analysis. A simulation study is conducted in Section 5 to quantify the systemic risk and diversification effect. Section 6 presents a sensitivity analysis. The last section concludes this study and ends with a discussion about future work.

2 Literature Review

Miranda and Glauber (1997) propose that the systemic risk problem may be a serious obstacle for an independent private crop insurance market to emerge. They investigate the ten largest U.S. crop insurers and indicate that the insurers face portfolio risks from 22 to 49 times larger than if the risks were independent. They conclude that systemic risk poses a pervasive problem for private crop insurance as the cost of maintaining adequate reserves to cover the large systemic risk is prohibitively high, unless efficient instruments transferring the systemic risk exist.

Several studies have attempted to identify the degree of the systemic risk at different regional levels and the possibility of spatial diversification of the systemic risk. Goodwin

(2001) evaluates the magnitude of spatial correlation of county-level corn yields for three major corn producing states (Illinois, Indiana, and Iowa). The author calculates the Pearson correlation coefficients along with a cubic regression function against the great-circle distance between each pair of counties. This work demonstrates that although the spatial correlation of corn yields is apparent, it does decay to a considerable degree as distance increases, suggesting the systemic risk could be diversified in a relatively large area. However, the author also mentions that the state-dependent feature of yield variables may affect the diversification effects.

Wang and Zhang (2003) apply a spatial statistics approach to investigate the extent of correlations of the county-level yield losses for three major crops (corn, soybeans, and wheat) in the US. They indicate that the correlations of yield losses for each crop fade out when the lag distance increases and point out that the systemic risk can be diversified if the risk pool is large enough. Therefore, the authors suggest that the risk pooling is effective and a private crop insurance market is possible in the U.S.

While these studies imply that it might be possible to diversify the systemic risk by increasing the trading area, they consider the systemic risk for only one crop within certain regions but not the risk across a number of crops. By contrast, Turvey, Nayak, and Sparling (1999) develop a theoretical model to evaluate portfolio reinsurance risks associated with agriculture. Eight crops in Ontario are selected to compose a portfolio, which implicitly accounts for the systemic risk across multiple crops. The joint distribution of crop yields is estimated by the mean and standard deviations of the sample over the years from 1985 to 1997. From the simulation results for reinsurance risks, the authors conclude that the systemic reinsurance risk is sensitive to the portfolio composition, but they do not examine whether the systemic risk can be diversified across crops or not.

Taking these early studies as a whole, it seems that they are all based on the assumption that the crop yields are linearly correlated. However, Goodwin (2001) indicates that the extent to which the yields are correlated may be state-dependent though his work is still based on linear correlations. He examines the spatial correlations of yields in normal years and in extreme weather years. This study shows that the spatial correlations decay as distance increases, but showing different patterns for the two cases. In normal years a faster decay is observed while in extreme weather years the decay is much slower, which means that the correlation of yields seems to be stronger in extreme years. Therefore, the author concludes that assuming the spatial correlations are linear may understate the magnitude of the systemic risk during extreme weather events.

Given the state-dependent feature of the correlations of crop yields, copulas can be considered as an effective alternative approach to modeling the correlations because of its greater flexibility in representing multivariate correlation structures. Copula approach was first introduced to empirical finance with regard to risk management by Embrechts, McNeil, and Straumann (1999). Recently, copulas have become a standard tool in finance and risk analysis. Widespread applications of copula approach have also been done in the field of insurance and actuarial mathematics (Junker and May, 2005; Trivedi and Zimmer, 2007).

With its increasing popularity in finance and insurance, copula approach has been adopted by recent researches on agriculture and crop insurance. Vedenov (2008) applies Gaussian and kernel copulas to model the correlation structure between the county-level and farm-level yields for Iowa corn. To illustrate the flexibility of copula functions, four different distributions (normal, gamma, Weibul, and nonparametric kernel density) are fitted to model marginal distributions of yields. From the estimation results, the correlation structures between the yields at the two different aggregation levels are not constant for different real-

izations of the yields, which can be effectively represented by applying copula approach to determine the correlation structure.

Zhu, Ghosh, and Goodwin (2008) use Gaussian and t copulas to define the joint yield and price risk of corn and soybeans for one county in Iowa. The authors compare these two types of copula functions by calculating the Akaike Information Criterion (AIC) and the log-likelihood value, obtaining a better goodness-of-fit for the t copula than the Gaussian copula. This implies tail dependence in the joint distribution of the yield and price risk as the t copula exhibits tail dependence while the Gaussian copula does not. By simulating values of prices and yields of corn and soybeans from the proposed copulas, the authors conclude that pooling risks across crops into a single insurance portfolio is superior to crop-specific insurance.

Several studies also make attempts to apply Archimedean copulas to investigate the correlation structure of multi-dimensional variables in the context of agricultural insurance. Xu *et al.* (2010) use three copulas from the Archimedean copula family to determine the level of the spatial correlations of different weather indices across regions at different aggregation levels in Germany. This work applies respective estimates of different copula functions to simulate the expected payoffs and buffer loads for weather-based insurance in different trading areas. The authors conclude that the possibility to reduce the systemic risk by increasing the trading area is limited for weather-based insurance in Germany.

Okhrin, Odening, and Xu (2012) argue that the use of exchangeable Archimedean copulas may cause some problems when estimating a high-dimensional correlation structure. They point out that exchangeable Archimedean copulas model the whole correlation structure with only one parameter. Consequently, in the case of high dimensionality, they may lead to large estimation errors and the substructure of the correlations is hidden. In addition, exchange-

able Archimedean copulas implicitly assume the exchangeability of the order of the marginal distributions within the copula functions. This implies that the permutation of the copula is symmetric, which is very restrictive for many applications, especially under high-dimensional cases. In view of these limitations, the authors employ the hierarchical Archimedean copulas (HAC) to explore the possibility of spatial diversification of the systemic weather risk in China. With the HAC estimates for the correlations of the weather indices for 17 regions in China, the authors simulate the buffer loads for a hypothetical weather-based insurance and reveals a significant spatial diversification effect on the weather risks. However, the authors indicate that despite the considerable diversification effect, the risk premiums are still too high for a viable private index-based insurance.

3 Econometric Framework

This section describes the framework of the copula-based approach for multivariate yield risks modeling. To measure the systemic risk inherent in crop yields, it is essential to determine the correlation structure among the multivariate random yield variables. The copula model is a very useful tool to model the joint distribution of the potentially high dimensional random yields. The joint distribution of the yield variables, which is represented by copula functions, can capture the potentially complicated correlation structure among these variables.

3.1 Basic copulas

The copula was first introduced by Sklar (1959). Sklar's theorem states that if F is an arbitrary k -dimensional joint continuous distribution function, then the associated copula is unique and defined as a continuous function $C : [0, 1]^k \rightarrow [0, 1]$ which satisfies the equation

$$F(x_1, \dots, x_k) = C[F_1(x_1), \dots, F_k(x_k)], \quad x_1, \dots, x_k \in \mathbb{R}, \quad (1)$$

where $F_1(x_1), \dots, F_k(x_k)$ are the respective marginal distributions.

In this way, the joint distribution of x_1, \dots, x_k can be described by the marginal distributions F_i and the correlation structure captured by the copula C . Note that the copula function is flexible in the sense that the variables x_i can be modeled with any kind of marginal distributions. Marginal distributions and the copula together uniquely determine the joint distribution. In turn, if the marginal distributions are continuous, a unique copula exists corresponding to the joint distribution. That is,

$$C(u_1, \dots, u_k) = F[F_1^{-1}(u_1), \dots, F_k^{-1}(u_k)], \quad u_1, \dots, u_k \in [0, 1], \quad (2)$$

where $F_1^{-1}(\cdot), \dots, F_k^{-1}(\cdot)$ are the corresponding quantile functions. Therefore, the copula can be defined as an arbitrary multivariate distribution on $[0, 1]^k$ with all marginal distributions being uniform.

Let c denote the density function of the copula C . Then c can be described as

$$c(u_1, \dots, u_k) = \frac{\partial^k C(u_1, \dots, u_k)}{\partial u_1 \cdots \partial u_k}, \quad (3)$$

The corresponding joint density function of x_1, \dots, x_k can then be written as

$$f(x_1, \dots, x_k) = c [F_1(x_1), \dots, F_k(x_k)] \prod_{i=1}^k f_i(x_i), \quad (4)$$

where $f_1(x_1), \dots, f_k(x_k)$ are marginal density functions.

Copula families are generally composed of parametric and nonparametric copulas. Empirical studies mainly apply the parametric copula because of its superiority in simulations. There are a large number of different parametric copula families. The most frequently used parametric copulas are the Gaussian (normal) copula and the Archimedean copulas, which imply different kinds of correlation structures.

The Gaussian copula takes the form of

$$C^N(u_1, \dots, u_k | \Sigma) = \Phi_\Sigma [\Phi^{-1}(u_1), \dots, \Phi^{-1}(u_k)], \quad (5)$$

where Φ_Σ is a k -dimensional normal distribution with zero mean and correlation matrix Σ , and Φ^{-1} is the inverse distribution function of the standard normal distribution. When modeling correlations, the Gaussian copula assumes linear correlations and implies no tail dependence.

In contrast to the Gaussian copula, the Archimedean copulas can be applied to model joint distributions with tail dependence. The general structure of a multivariate Archimedean copula has the following form

$$C(u_1, \dots, u_k) = \phi^{-1} [\phi(u_1) + \dots + \phi(u_k)], \quad (6)$$

where ϕ is called the generator of the Archimedean copula and ϕ^{-1} is inverse of ϕ .

Two widely used multivariate Archimedean copulas are rather effective in modeling asymmetric tail dependence. One is the Gumbel copula that implies stronger upper tail depen-

dence and relatively weaker lower tail dependence. The cumulative density function (cdf) of the Gumbel copula can be written as

$$C^G(u_1, \dots, u_k | \theta) = \exp \left\{ - \left[(-\log u_1)^\theta + \dots + (-\log u_k)^\theta \right]^{1/\theta} \right\}, \quad 1 \leq \theta < \infty, \quad (7)$$

where θ denotes the Gumbel copula parameter. $\theta = 1$ represents independence and $\theta \rightarrow \infty$ represents perfect positive dependence. The generator and its inverse of the Gumbel copula are

$$\phi(u, \theta) = (-\log u)^\theta, \quad (8)$$

$$\phi^{-1}(u, \theta) = \exp(-u^{1/\theta}).$$

The Gumbel copula can also be used to model strong lower-tail dependence by applying the survival Gumbel copula. Following Sklar's theorem, the survival copula is defined as follows:

Let \bar{F} be a k -dimensional joint survival function with continuous marginal survival functions $\bar{F}_1, \dots, \bar{F}_k$. Then a unique survival copula \bar{C} exists such that

$$\bar{F}(x_1, \dots, x_k) = \bar{C} \{ \bar{F}_1(x_1), \dots, \bar{F}_k(x_k) \}, \quad x_1, \dots, x_k \in \mathbb{R}. \quad (9)$$

The other is the Clayton copula which implies stronger lower tail dependence and relatively weaker upper tail dependence. The cumulative density function (cdf) of the Clayton copula is given by

$$C^C(u_1, \dots, u_k | \theta) = [u_1^{-\theta} + \dots + u_k^{-\theta} - (k-1)]^{-1/\theta}, \quad 0 < \theta < \infty, \quad (10)$$

where θ denotes the Clayton copula parameter. $\theta \rightarrow 0$ represents independence and $\theta \rightarrow \infty$ represents perfect positive dependence. The generator and its inverse of the Clayton copula are

$$\phi(u, \theta) = \frac{1}{\theta} (u^{-\theta} - 1), \quad (11)$$

$$\phi^{-1}(u, \theta) = (\theta u + 1)^{-1/\theta}.$$

Another commonly used Archimedean copula is the Frank copula. While the Gumbel and Clayton copula only allow for positive dependence, the Frank copula can exhibit both positive and negative dependence. However, the Frank copula does not imply any tail dependence. The cumulative density function (cdf) of the Frank copula is given by

$$C^F(u_1, \dots, u_k | \theta) = -\frac{1}{\theta} \log \left[1 + \frac{\prod_{i=1}^k (e^{-\theta u_i} - 1)}{(e^{-\theta} - 1)^{k-1}} \right], \quad -\infty < \theta < \infty, \theta \neq 0, \quad (12)$$

where θ denotes the Frank copula parameter. $\theta \rightarrow -\infty$ represents perfect negative dependence, $\theta \rightarrow 0$ represents independence, and $\theta \rightarrow \infty$ represents perfect positive dependence. The generator and its inverse of the Frank copula are

$$\phi(u, \theta) = -\log \left(\frac{\exp(-\theta u) - 1}{\exp(-\theta) - 1} \right), \quad (13)$$

$$\phi^{-1}(u, \theta) = -\frac{1}{\theta} [\log 1 + \exp(-u)(\exp(-\theta) - 1)].$$

3.2 Hierarchical Kendall copulas

The basic copulas have become popular in modeling correlations of different kinds of structure. However, they are only effective for low dimensional situations, such as pair-wise cases.

In higher dimensions, the multivariate models for a basic copula often turn out to be inflexible and restrictive. For example, the multivariate Gaussian copula requires the specification of the correlation matrix that can involve a large number of parameters in high dimensional cases ($n(n - 1)/2$ parameters for n -dimensional cases). Furthermore, the Gaussian copula can only model symmetric dependence, not accounting for any kind of tail dependence. The Archimedean copulas, while allowing for asymmetric tail dependence, imply symmetry of the permutation of variables within the copula function. That is, the order of the marginal distributions u_i is exchangeable. This restriction seems to be very strict and implausible. In addition, the Archimedean copulas represent the multivariate correlation structure with only one single parameter, which implies no substructure of the correlations and may lead to large estimation errors.

Recent research has developed some advanced copula models to deal with correlations involved with high dimensional variables. Such copula models include vine copulas and hierarchical Archimedean copulas. A vine copula is built by decomposing the joint multivariate density into a product of pair-copulas, which allows for considerable flexibility in high dimensions. However, lack of parsimony can be a problem for the vine copula in very high dimensions. The extreme numbers of parameters in such cases ($n(n - 1)/2$ parameters for n -dimensional cases) may severely affect the availability of the vine copula from a computational viewpoint. Hierarchical Archimedean copulas are much more parsimonious in terms of the number of the parameters ($n - 1$ parameters for n -dimensional cases). However, they are more restrictive at the same time in that the building blocks are restricted to the class of Archimedean copulas, which is too strict and not appropriate in some applications.

The hierarchical Kendall copulas, a recently developed innovation, have achieved both flexibility and parsimony when modeling the joint distribution of high dimensional variables.

A hierarchical Kendall copula is built by a hierarchy of copulas, which ensures that it is a parsimonious model. However, in contrast to hierarchical Archimedean copulas, the choice of the copula at each hierarchical level is not limited to any class of copulas. In other words, the building blocks can be copulas of arbitrary types. As hierarchical Kendall copulas share the property of flexibility as well as parsimony, they have been considered as an extremely useful tool for modeling the potentially complex correlation structure among a large number of variables. Therefore, in this study we focus on the hierarchical Kendall copula model and apply it to reveal the correlation structure of the high dimensional yield variables. More details about this copula model are introduced as follows.

3.2.1 Kendall distribution functions

An important component of hierarchical Kendall copulas is the Kendall distribution function. Following Genest and Rivest (1993) and Brechmann (2013), for $U := (U_1, \dots, U_d)' \sim C$, where C is a d -dimensional copula, the Kendall distribution function $K^{(d)}$ is defined as

$$K^{(d)}(t) := P(C(U) \leq t), \quad t \in [0, 1]. \quad (14)$$

For $t \in [0, 1]$, it holds that $t \leq K^{(d)}(t) \leq 1$ for $t \in [0, 1]$ and $K^{(d)}(0-) = 0$. From the definition, the Kendall distribution function is the univariate distribution function of the random variable $Z := C(U)$. Thus it holds that $K^{(d)}(Z) \sim U(0, 1)$.

It is in general complicated to derive the Kendall distribution function in explicit form for a given copula. A recursive formula is given by Imlahi et al. (1999):

$$K^{(d)}(t) = K^{(d-1)}(t) + \int_t^1 \int_{C_{u_1}^{-1}(t)}^1 \cdots \int_{C_{u_1, \dots, u_{d-2}}^{-1}(t)}^1 \int_0^{C_{u_1, \dots, u_{d-1}}^{-1}(t)} c(u_1, \dots, u_d) du_d \dots du_1, \quad (15)$$

where $K^{(d)}$ denotes the Kendall distribution function of the d -dimensional copula C with the density c and $K^{(d-1)}$ denotes the Kendall distribution function of the $(d-1)$ -dimensional margin of the first $d-1$ variables. The formula is also involved with the inverse of the copula quantile function C_{u_1, \dots, u_r}^{-1} which is defined as

$$C(u_1, \dots, u_r, C_{u_1, \dots, u_r}^{-1}(z), 1, \dots, 1) = z, \quad (16)$$

for $r = 1, \dots, d-1$, and

$$C_{\emptyset}^{-1}(z) := z,$$

for $z \in (0, 1)$.

The usually high-dimensional integration and nonavailability of the copula quantile function in explicit form in equation (15) often make it not possible to determine the Kendall distribution function explicitly. One exception is the Archimedean copulas. Barbe et al. (1996) derived the Kendall distribution function for a d -dimensional Archimedean copula with generator φ as

$$K^{(d)}(t) = t + \sum_{i=1}^{d-1} \frac{(-1)^i}{i!} \varphi(t)^i (\varphi^{-1})^{(i)}(\varphi(t)), \quad (17)$$

for $t \in (0, 1]$, where $(\varphi^{-1})^{(i)}(\cdot)$ denotes the i -th derivative of φ^{-1} .

3.2.2 Definition of hierarchical Kendall copulas

Brechmann (2013) defined the Hierarchical Kendall copula from the Kendall distribution function as following:

“Let $u_1, \dots, u_n \sim U(0, 1)$ and let C_0, C_1, \dots, C_d be copulas of dimensions d, n_1, \dots, n_d , respectively, where $n_i \geq 1, i = 1, \dots, d$, and $n = \sum_{i=1}^d n_i$. Let K_1, \dots, K_d denote the Kendall distribution functions corresponding to C_1, \dots, C_d , respectively. Define $m_i = \sum_{j=1}^i n_j$ for $i = 1, \dots, d$, and $m_0 = 0$ as well as $U_i := (u_{m_{i-1}+1}, \dots, u_{m_i})'$ and $V_i := K_i(C_i(U_i))$ for $i = 1, \dots, d$. Under the assumptions that

$A_1 : U_1, \dots, U_d$ are mutually independent conditionally on $(V_1, \dots, V_d)'$, and

$A_2 : \text{the conditional distribution of } U_i \mid (V_1, \dots, V_d)'$ is the same as the conditional distribution of $U_i \mid V_i$ for all $i = 1, \dots, d$, that is, $F_{U_i|V_1, \dots, V_d} = F_{U_i|V_i} \forall i \in \{1, \dots, d\}$,

the random vector $(u_1, \dots, u_n)'$ is said to be distributed according to the hierarchical Kendall copula C_K with nesting copula C_0 and cluster copulas C_1, \dots, C_d if

(i) $U_i \sim C_i \forall i \in \{1, \dots, d\}$,

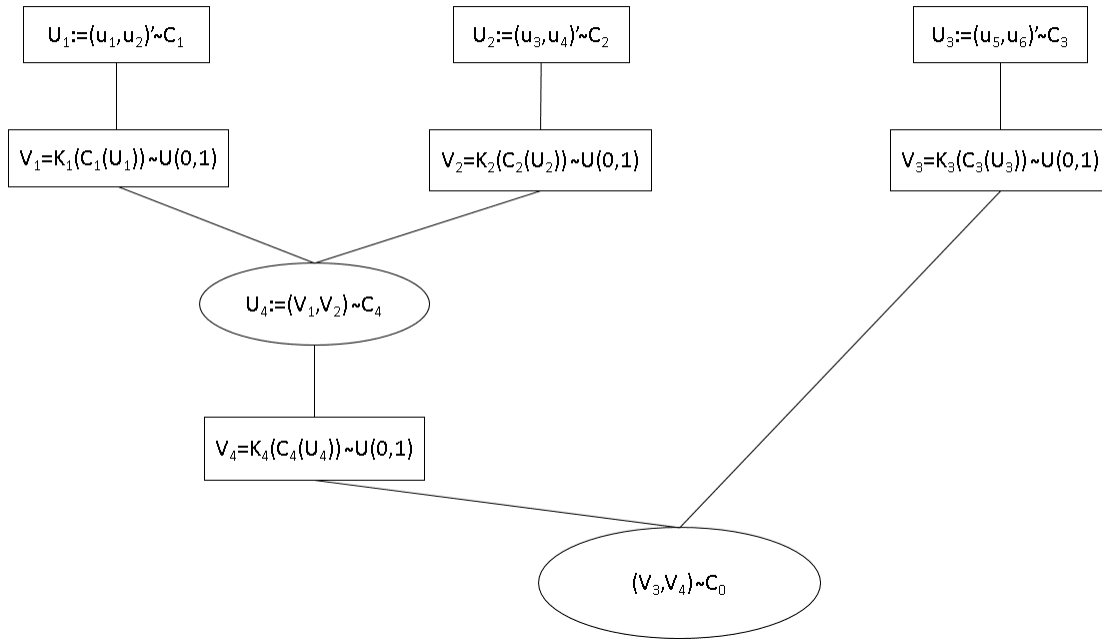
(ii) $(V_1, \dots, V_d)' \sim C_0$.”

As Brechmann (2013) notes, the two assumptions can be interpreted as that, conditional on the information of the nesting variables V_1, \dots, V_d , the clusters U_1, \dots, U_d are independent of each other and also independent of other nesting variables. That is, while the correlations within each cluster U_i are explained by the joint behavior of the random variables $u_{m_{i-1}+1}, \dots, u_{m_i}$ for $i = 1, \dots, d$, the correlations among the clusters U_1, \dots, U_d are explained through the joint behavior of the unobserved factors V_1, \dots, V_d , each of which has a uniform distribution because $C_i(U_i) \sim K_i$ for all $i = 1, \dots, d$.

The above definition from Brechmann (2013) has indicated how the two-level hierarchical

Kendall copula is built. The construction can be extended to an arbitrary number of levels. Figure 1 is an illustration of a simple three-level hierarchical Kendall copula. For more detailed descriptions of the general k -level hierarchical Kendall copulas, reference can be found in Brechmann (2013).

Figure 1: **Illustration of a three-level hierarchical Kendall copula**



4 Empirical Application

The empirical application in this section is a preliminary study intended to apply a copula-based approach to examine the effectiveness of diversifying systemic risk across two countries and five crops. The study focuses on two large agricultural producing countries, the United States and China. Five major crops in these two countries are selected, which are corn, cotton, rice, soybeans, and wheat. The joint distribution of the multivariate yield variables is modeled by a hierarchical Kendall copula model.

4.1 Study area and data

The study area includes ten major producing states in the US and ten major producing provinces in China for each crop except that only six states are available for the rice in the US. So the US crop yield data include 46-dimensional state-level historical yields, and the China crop yield data include 50-dimensional province-level historical yields. The US crop yield data, covering 44-year period from 1970 to 2013, were taken from the USDA's National Agricultural Statistics Service (NASS) databases. The China crop yield data, covering only 31-year period spanning from 1979 to 2009 due to the lack of data for other years, were taken from China Statistical Yearbook.

4.2 Detrending the yield data

Technological advancements in crop production result in an upward trend in crop yields over time. To adjust for this trend in the historical yield data, an initial step is to apply a detrending process. The observed yield data were detrended by estimating the equation:

$$y_t = \beta_0 + \beta_1 t + \epsilon_t, \quad (18)$$

where $t = 1970, \dots, 2013$ for each of the US state-level yield data, $t = 1979, \dots, 2009$ for each of the China province-level yield data. The corresponding yield trends were calculated as the predicted yields from the above regression:

$$\hat{y}_t = \beta_0 + \beta_1 t, \quad (19)$$

Detrended yields to 2009 equivalents were generated as:

$$y_t^{det} = y_t \frac{\hat{y}_{2009}}{\hat{y}_t}. \quad (20)$$

where $t = 1970, \dots, 2013$ for the US, $t = 1979, \dots, 2009$ for China. All the yield data used in the remainder of this study is composed of the detrended 2009-equivalent yields.

4.3 Estimation

A hierarchical Kendall copula model was fitted to the detrended yield data. We adopted a commonly used two-step procedure to estimate the parameters of the copula model. This estimation method is called inference for margins (IFM) (Joe and Xu, 1996). In the IFM method, the parameters of the marginal distributions are estimated first. Next, the copula

parameters are determined given the estimated margins \hat{F}_{X_i} , $i = 1, \dots, n$. That is, we estimate the parameters of the hierarchical Kendall copula based on the uniform variables derived as $\hat{u}_{ji} = \hat{F}_{X_i}(x_{ji})$, $j = 1, \dots, t$, $i = 1, \dots, n$. As demonstrated by Joe (2005), the IFM method provides consistent estimators of copula parameters, and almost loses no efficiency but being computationally much more attractive compared to joint estimation of the parameters of margins and copula functions.

4.3.1 Estimation of the marginal distributions

At the first step, we independently estimate the marginal distributions for each of the state (province)-level yield variables. As crop yields tend to be negatively skewed (Ramirez, 1997), the beta distribution has been frequently used in many studies to capture the potential skewness for crop yields (Babcock and Hennessy, 1996; Nelson and Preckel, 1989). The beta density function of yield variable y can be written as:

$$Beta(y \mid \alpha, \beta, y_{min}, y_{max}) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} \frac{(y - y_{min})^{\alpha-1} (y_{max} - y)^{\beta-1}}{y_{max}^{\alpha+\beta-1}}, \quad y_{min} \leq y \leq y_{max}, \quad (21)$$

where $\Gamma(\cdot)$ denotes the gamma function, y_{min} and y_{max} are parameters that denote the lower and upper limit of the feasible range for y respectively, and α and β are shape parameters.

We fitted a separate beta distribution to the historical data of each state (province)-level yield variable y_i . For each y_i , the lower and upper limit were set as

$$y_{min_i} = 0,$$

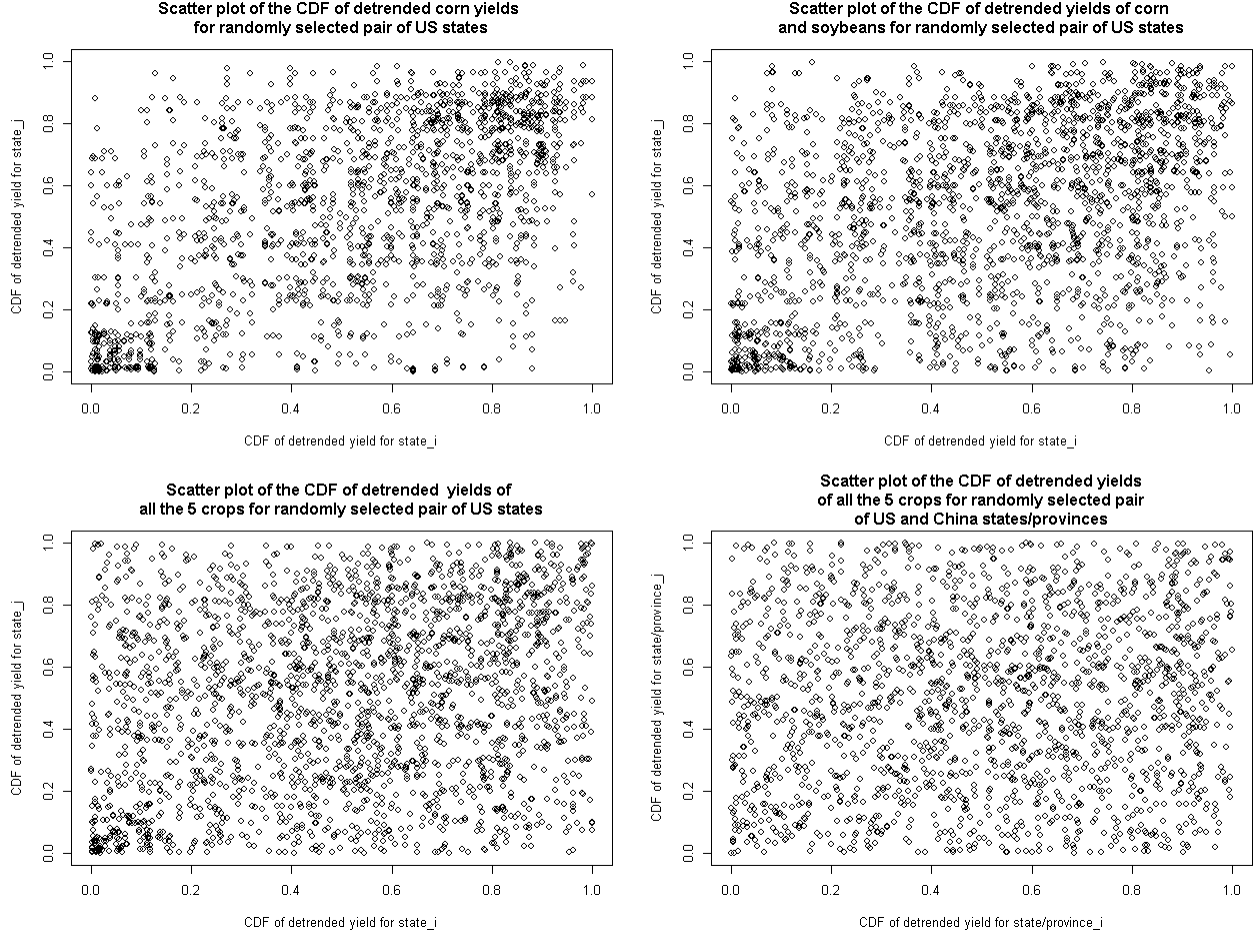
$$y_{max_i} = y_i^U + 1.5\sigma_i,$$

where y_i^U denotes the maximum observed yield of y_i , and σ_i is the sample standard deviation of y_i . The shape parameters α_i and β_i were estimated by the maximum likelihood method.

4.3.2 Copula model selection

The second step of the IFM method involves the estimation of copula functions, which is based on the uniform variables $u_i = \hat{F}_i(y_i)$, where \hat{F}_i is the cumulative distribution function (cdf) of the estimated beta distribution for yield variable y_i in the first step. Before formal copula modeling, we show some visual patterns of the uniform variables regarding the correlations among them (Figure 2). From the plots, we observe that while risk-pooling across different types of crops within the US mitigates the magnitude of the systemic nature (lower-tail dependence) of yield risks, combining the two countries (the US and China) seems to eliminate the systemic nature of risks, leaving no tail-dependence at all.

Figure 2: Scatter plots of the CDF of randomly paired detrended yields



Given the two geographic clusters (the US and China) and the observations above, hierarchical Kendall copulas are very suitable for modeling the correlations among these variables. The nesting copula is chosen to represent the correlation between these two clusters. Within each cluster, the correlation structure is built up by aggregating a group of variables by one cluster copula at each hierarchical level. To reveal as much relevant information as possible on the correlation structure as well as the substructure of the correlations, bivariate copulas were used as building blocks at all hierarchical levels. That is, we joined two variables

together each time. As crop yields tend to be more highly correlated during widespread extreme weather events within a certain area (Goodwin, 2001), resulting large systemic yield losses, cluster copulas should be selected to be able to account for this strong lower-tail dependence. In this study, we tried two different kinds of copulas that are effective in modeling lower-tail dependence, the bivariate survival Gumbel copula and the bivariate Clayton copula, as the building blocks within each cluster. By contrast, we assumed no tail dependence for the correlation between the two geographic clusters because of the long distance and so almost uncorrelated weather patterns between the US and China. Therefore, the Gaussian copula and the Frank copula, both of which imply tail independence, were chosen as our alternatives for the nesting copula.

To specify the whole hierarchical structure within each cluster, we adopt hierarchical clustering mechanism with an appropriate metric to measure the distance between (groups of) variables. Variables with shorter distance are aggregated at lower hierarchical level. The distance between two variables is determined by the level of correlation between them. The higher the correlation, the shorter the distance. As the copula parameter is increasing with the associated (rank) correlation level for both survival Gumbel and Clayton copulas, the distance can be measured by the value of the copula parameters. Specifically, following Okhrin and Ristig (2012), the whole hierarchical structure was constructed in the following recursive way. At the lowest level, a bivariate cluster copula was fitted to every possible couple of the variables u_i by the maximum likelihood method. We selected the couple of variables (denoted as \bar{u}_1 and \bar{u}_2) with the highest level of correlation and denoted the estimated bivariate copula and its parameter as C_1 and $\hat{\theta}_1$, respectively. Then we introduced a pseudo-variable $Z_1 = C_1(\bar{u}_1, \bar{u}_2; \hat{\theta}_1)$. At the next step, the remaining variables and the pseudo-variable composed a new set. We proceeded in the same way considering this new

set of variables. We chose the most highly correlated couple of variables from this new set and obtained another pseudo-variable from the respective estimated bivariate copula. We repeated this procedure until the whole hierarchical structure was determined.

4.3.3 Estimation of the copula parameters in the presence of estimation risk

The parameters of a Hierarchical Kendall copula with specified structure can be estimated sequentially with the following algorithm given by Brechmann (2013). For a two-level hierarchical Kendall copula C_K with nesting copula C_0 and cluster copulas C_1, \dots, C_d , let $(u_{j,1}, \dots, u_{j,n})'_{j=1, \dots, t}$ be a sample of the hierarchical Kendall copula C_K , where C_0, C_1, \dots, C_d are copulas of dimensions d, n_1, \dots, n_d , respectively. Define $m_i = \sum_{l=1}^i n_l$ for $i = 1, \dots, d$, and $m_0 = 0$. The estimates of the parameters $\theta_0, \theta_1, \dots, \theta_d$ of the copulas C_0, C_1, \dots, C_d respectively are obtained by

(i) estimating θ_i based on $(u_{j,m_{i-1}+1}, \dots, u_{j,m_i})'_{j=1, \dots, t}$ by the maximum likelihood method, for $i = 1, \dots, d$, and

(ii) estimating θ_0 based on the pseudo observations

$$\hat{v}_{j,i} := K_i(C_i(u_{j,m_{i-1}+1}, \dots, u_{j,m_i}; \hat{\theta}_i); \hat{\theta}_i), \quad i = 1, \dots, d, \quad j = 1, \dots, t, \quad (22)$$

by the maximum likelihood method. This sequential procedure can be generalized to estimate k -level hierarchical Kendall copulas by proceeding with more steps.

However, the above estimation procedure, which involves plugging-in the estimates $\hat{\theta}_i$ from the lowest level to the highest level step by step, ignores estimation risk in general. Given the high dimensions of the hierarchical Kendall copula applied in this study, ignoring

estimation risk may lead to accumulating estimation errors at each sequential estimation step, especially when we have only a short time series of available yield observations.

To increase efficiency of the estimation, we propose a different sequential approach that takes estimation risk into account. This approach is based on Bayes' criterion. With the same specification as the previous algorithm, we obtain estimates of the parameters $\theta_0, \theta_1, \dots, \theta_d$ by

(i) estimating θ_i based on $(u_{j,m_{i-1}+1}, \dots, u_{j,m_i})'_{j=1,\dots,t}$ by the Bayesian method and obtaining the posterior density $p(\theta_i | u_{j,m_{i-1}+1}, \dots, u_{j,m_i})$ for θ_i , for $i = 1, \dots, d$, and

(ii) estimating θ_0 based on the pseudo observations

$$\hat{v}_{j,i} := \int_{\Theta} K_i(C_i(u_{j,m_{i-1}+1}, \dots, u_{j,m_i}; \theta_i); \theta_i) p(\theta_i | u_{j,m_{i-1}+1}, \dots, u_{j,m_i}) d\theta_i, \quad i = 1, \dots, d, \quad j = 1, \dots, t, \quad (23)$$

by the Bayesian method. This procedure can also be generalized to estimate k -level hierarchical Kendall copulas.

We used the proposed procedure to estimate the hierarchical Kendall copula based on the yield data set. The cluster copulas (building blocks) and nesting copula are all bivariate copulas. The posterior density for the bivariate copula parameter θ conditional on a sample of two variables (u_1, u_2) , according to Bayes' theorem, is the product of the likelihood function $f(u_1, u_2 | \theta)$ and the prior distribution $\pi(\theta)$ normalized by an appropriate constant:

$$p(\theta | u_1, u_2) = \frac{f(u_1, u_2 | \theta) \pi(\theta)}{\int f(u_1, u_2 | \theta) \pi(\theta) d\theta}. \quad (24)$$

Given a sample $(u_{j,1}, u_{j,2})'_{j=1,\dots,t}$ of the variables (u_1, u_2) , the likelihood function is given by

$$f(u_1, u_2 | \theta) = \prod_{t=1}^T f(u_{1t}, u_{2t} | \theta) = \prod_{t=1}^T c(u_{1t}, u_{2t} | \theta), \quad (25)$$

where $c(\cdot, \cdot)$ is the copula density function derived by taking derivatives of the copula function C :

$$c(u_1, u_2 | \theta) = \frac{\partial^2 C(u_1, u_2 | \theta)}{\partial u_1 \partial u_2}. \quad (26)$$

For the bivariate survival Gumbel copula, the density function is derived as

$$\bar{c}^G(\bar{u}_1, \bar{u}_2 | \theta) = \frac{\left[(-\ln \bar{u}_1)^\theta + (-\ln \bar{u}_2)^\theta\right]^{\frac{1}{\theta}} + \theta - 1}{\bar{u}_1 \bar{u}_2} \exp \left\{ - \left[(-\ln \bar{u}_1)^\theta + (-\ln \bar{u}_2)^\theta\right]^{1/\theta} \right\} \\ \left[(-\ln \bar{u}_1)^\theta + (-\ln \bar{u}_2)^\theta\right]^{\frac{1}{\theta}-2} (-\ln \bar{u}_1)^{\theta-1} (-\ln \bar{u}_2)^{\theta-1}, \quad (27)$$

where $\bar{u}_i = 1 - u_i$ for $i = 1, 2$ are “survival variables”.

For the bivariate Clayton copula, the density function is derived as

$$c^C(u_1, u_2) = (\theta + 1) (u_1^{-\theta} + u_2^{-\theta} - 1)^{-\frac{1}{\theta}-2} u_1^{-\theta-1} u_2^{-\theta-1}. \quad (28)$$

For the bivariate Frank copula, the density function is derived as

$$c^F(u_1, u_2) = - \frac{\theta \exp(-\theta u_1) \exp(-\theta u_2) [\exp(-\theta) - 1]}{\{[\exp(-\theta u_1) - 1][\exp(-\theta u_2) - 1] + \exp(-\theta) - 1\}^2}. \quad (29)$$

For the bivariate Gaussian copula, the density function is derived as

$$c^N(v_1, v_2 | \Sigma) = |\Sigma|^{-\frac{1}{2}} \exp \left\{ - \frac{[\Phi^{-1}(v_1), \Phi^{-1}(v_2)]' (\Sigma^{-1} - I) [\Phi^{-1}(v_1), \Phi^{-1}(v_2)]}{2} \right\}, \quad (30)$$

where Φ_{Σ} is a 2-dimensional normal distribution with zero mean and correlation matrix Σ , Φ^{-1} is the inverse distribution function of the standard normal distribution, and I is the identity matrix.

To prevent any effect of prior information on the estimation results, we selected a uniform distribution as the prior for all the bivariate Archimedean copula parameter θ 's. The bounds were selected so that a wide range of the support for θ is covered:

$$Gumbel : \theta \sim U(1, 100)$$

$$Clayton : \theta \sim U(0, 100)$$

$$Frank : \theta \sim U(-1000, 1000).$$

Following Smith (2011), we selected the prior for the correlation matrix of the bivariate Gaussian copula based on a Cholesky decomposition. The correlation matrix Σ can be decomposed as

$$\Sigma = \text{diag}(\Omega)^{-\frac{1}{2}} \Omega \text{diag}(\Omega)^{-\frac{1}{2}}, \quad (31)$$

where Ω is a positive definite matrix. And the matrix Ω can be decomposed as

$$\Omega = RR', \quad (32)$$

where $R = \{r_{i,j}\}$ is a lower triangular Cholesky factor. To ensure the decompositions are unique, we set $r_{i,i} = 1$, for $i = 1, 2$. Then the correlation matrix Σ can be written as

$$\Sigma = \begin{bmatrix} 1 & \frac{r_{2,1}}{\sqrt{1+r_{2,1}^2}} \\ \frac{r_{2,1}}{\sqrt{1+r_{2,1}^2}} & 1 \end{bmatrix}. \quad (33)$$

We assigned a non-informative prior distribution of $r_{2,1}$ as $r_{2,1} \sim U(-100, 100)$ to cover a wide range of the support for the correlation matrix.

We used the above Bayesian models to estimate the copula parameters by simulating the posterior distribution by Markov Chain Monte Carlo (MCMC) methods. MCMC techniques avoid explicit evaluation of the rather complex integrals that is hardly possible in this study. For each of the copula Bayesian models, three chains of 600,000 iterations were run from different initial values. The first 300,000 iterations were discarded as a burn-in period. To reduce the autocorrelation, we saved every 30th iteration of each chain. The adequate convergence was confirmed by both the Monte Carlo error (<0.001) and the Gelman-Rubin (1992) test (the potential scale reduction factor is 1) for all parameters.

Note that within each of the two clusters (the US and China), the Bayesian inference for the parameters of the cluster copulas was based on all the available data. For the US, it was based on the historical yield data covering 44 years (1970-2013), and resulted in the pseudo observations $\hat{v}_{1970,us}, \dots, \hat{v}_{2013,us}$. For China, it was based on the historical yield data covering 31 years (1979-2009), and resulted in the pseudo observations $\hat{v}_{1979,ch}, \dots, \hat{v}_{2009,ch}$. As we do not have the historical yield data for China before year 1979 or after year 2009, the Bayesian inference for the parameters of the nesting copula was based on part of the US pseudo observations and all the China pseudo observations: $\hat{v}_{1979,us}, \dots, \hat{v}_{2009,us}$ and $\hat{v}_{1979,ch}, \dots, \hat{v}_{2009,ch}$. This also demonstrates the flexibility of the hierarchical Kendall copula model. Although the information contained in the US data before 1979 or after 2009 cannot be used for the nesting copula, it still can be used for the cluster copulas withing the US. With such a copula model, all the relative information in the available data can be exploited instead of wasting some of it.

Kendall's correlation coefficient is used in this study as a measure of association between

the correlation level and the copula function, since it is indifferent to nonlinear monotonic transformations. Following Nelsen (2006), the Kendall's τ associated with Archimedean copulas can be expressed as

$$\begin{aligned} \text{Gumbel} : \tau &= 1 - \frac{1}{\theta} \\ \text{Clayton} : \tau &= \frac{\theta}{\theta + 2} \\ \text{Frank} : \tau &= 1 - \frac{4}{\theta} + \frac{4}{\theta^2 \int_0^\theta \frac{t}{\exp(t)-1} dt}. \end{aligned}$$

The Kendall's τ associated with the Gaussian copula can be written as

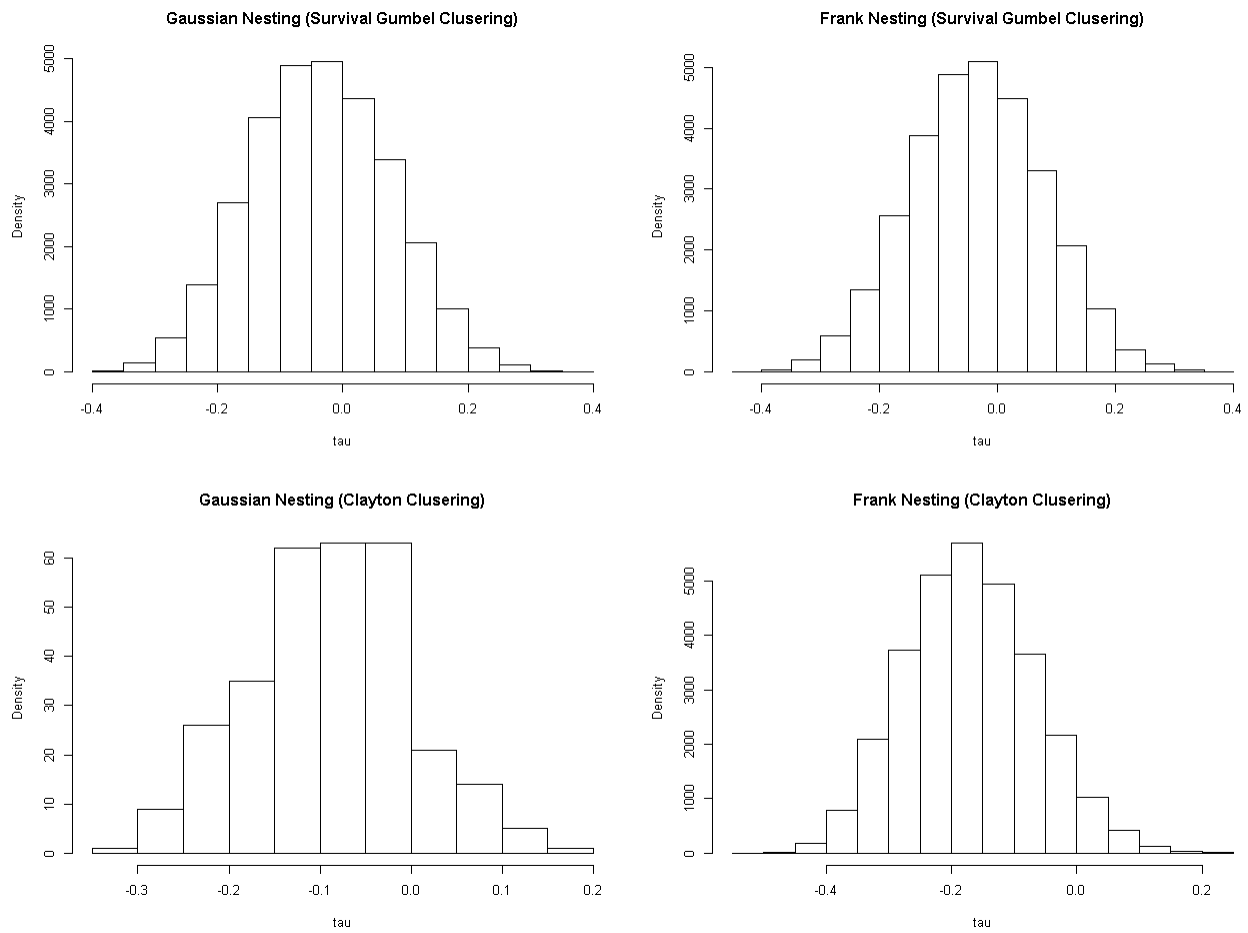
$$\tau(v_i, v_j) = \frac{2}{\pi} \arcsin(\epsilon_{ij}),$$

where ϵ_{ij} is the (i, j) th element of the correlation matrix Σ of the Gaussian copula. Within the U.S., the posterior mean of the Kendall's τ associated with the cluster copulas ranges from 0.111 to 0.698 as implied by survival Gumbel copula, and from 0.091 to 0.638 as implied by Clayton copula. Within China, the posterior mean of the Kendall's τ associated with the cluster copulas ranges from 0.072 to 0.613 as implied by survival Gumbel copula, and from 0.046 to 0.575 as implied by Clayton copula. This indicates that there are certain levels of positive correlations among the yield variables within each country.

The between-country correlations are calculated from the nesting copulas. The posterior distribution of the Kendall's τ associated with the nesting copulas was plotted in Figure 3. From the graph, the posterior mean and median are very close to zero under all models. In addition, with a posterior probability of 95%, the correlation is less than 0.151 (0.067) for Gaussian nesting with survival Gumbel (Clayton) clustering, and less than 0.152 (0.003) for Frank nesting with survival Gumbel (Clayton) clustering, all indicating that there is little

positive correlation of the crop yields between the US and China. Therefore, it seems quite possible that we could diversify the yield risks effectively across these two countries, which has been investigated in the following simulation study.

Figure 3: The Posterior Distribution of the Kendall's Correlation Coefficient for the Nesting Copula



5 Simulation Study

With the simulated copula parameters from the full posterior distribution and the estimated Beta margins, a sample of predicted yields can be generated. The predictive distribution of aggregated net insurance income at different coverage levels can then be obtained based on the predicted yields. The systemic risk associated with an insurance portfolio is assessed by the statistics of the predictive distribution of net insurance income.

According to Bayes' Theorem, the posterior predictive distribution of a new unknown observable \tilde{u} conditional on the observed data u is

$$p(\tilde{u} | u) = \int p(\tilde{u}, \theta | u) d\theta = \int p(\tilde{u} | \theta, u) p(\theta | u) d\theta = \int p(\tilde{u} | \theta) p(\theta | u) d\theta. \quad (34)$$

Recall that we have independently simulated a sample of 30000 values from the posterior distribution for each of the hierarchical Kendall copula parameters. So a sample of predicted vectors of the uniforms can be generated by simulating the s th vector of uniforms from the corresponding hierarchical Kendall copula with the s th simulated copula parameters, for $s = 1, \dots, 30000$.

Sampling from a given hierarchical Kendall copula can be conducted with the following algorithm provided by Brechmann (2013). To sample u_1, \dots, u_n from a two-level hierarchical Kendall copula C_K with nesting copula C_0 and cluster copulas C_1, \dots, C_d , we proceed by

- (i) Sample v_1, \dots, v_d from C_0 .
- (ii) Set $z_i := K_i^{-1}(v_i)$, where K_i^{-1} denotes the inverse of the Kendall distribution function K_i , for $i = 1, \dots, d$.
- (iii) Sample u_1, \dots, u_n from $(U_{m_{i-1}+1}, \dots, U_{m_i}) \mid C_i(U_{m_{i-1}+1}, \dots, U_{m_i}) = z_i$ for $i = 1, \dots, d$.

This procedure can be extended for sampling from a general k -level hierarchical Kendall copula by adding more iterations. Note that after sampling from the nesting copula C_0 , simulating from each of the cluster copulas involves sampling from a general conditional distribution $U \mid C(U) = z$. As the cluster copulas are all bivariate Archimedean copulas in this study, we are to sample from the conditional distribution $(U_1, U_2) \mid C(U_1, U_2) = z$. This can be accomplished by the conditional inverse method (Devroye, 1986), in which u_1 is sampled from the conditional distribution $U_1 \mid C(U_1, U_2) = z$ and u_2 is set to satisfy that $C(u_1, u_2) = z$. As demonstrated by Brechmann (2013), the conditional distribution function of $U_1 \mid C(U_1, 1) = z$ is derived as

$$F_1(u \mid C(U_1, U_2) = z) = \frac{\int_z^u g_1(u_1) du_1}{\int_z^1 g_1(u_1) du_1}, \quad u \in (z, 1), \quad (35)$$

where $g_1(u_1) = \int_{C_{u_1}^{-1}(z)}^1 \int_z^1 c(u_1, C_{u_1}^{-1}(z)) \frac{\partial C_{u_1}^{-1}(z)}{\partial z} du_1 du_2$. When C is an Archimedean copula with generator φ , $F_1(u \mid C(U_1, U_2) = z)$ can be written in closed form as

$$F_1(u \mid C(U_1, U_2) = z) = 1 - \frac{\varphi(u)}{\varphi(z)}, \quad u \in (z, 1). \quad (36)$$

Therefore, u_1 and u_2 can be sampled with the following algorithm:

- (i) Sample w from $U(0, 1)$.
- (ii) $u_1 = \varphi^{-1}((1 - w)\varphi(z))$.
- (iii) $u_2 = \varphi^{-1}(\varphi(z) - \varphi(u_1))$.

Given the simulated vectors of uniform variates from each copula model, predicted yields were obtained based on the estimated marginal distributions.

We consider an insurance portfolio composed of state (province)-level area yield insurance contracts at different aggregated levels. For each state (province)-level insurance contract i ,

the indemnity paid for one unit area takes the form of

$$I_i = p_i * \max[\lambda y_i^e - y_i, 0], \quad (37)$$

where y_i denotes the yield, $y_i^e = E(y_i)$ denotes the expected yield, λ denotes the coverage level, and p_i denotes the base price of the crop associated with insurance contract i , which represents the expected harvest time price at planting time. We consider two coverage levels 70% and 90% in this study. The base price is selected as the average price during the planting month of futures contracts expiring at harvest time. The actuarially fair premium π_i is equal to the expected indemnity payment:

$$\pi_i = E(I_i) = p_i * E\max[\lambda y_i^e - y_i, 0] = p_i * E[(\lambda y_i^e - y_i)I(y_i \leq \lambda y_i^e)], \quad (38)$$

where I is the indicator function. The net insurance loss of insurance contract i can then be written as

$$L_i = I_i - \pi_i. \quad (39)$$

The aggregated net insurance loss L associated with an insurance portfolio is therefore the weighted average of L_i :

$$L = \sum w_i L_i, \quad (40)$$

where w_i denotes the weight of insurance contract i , which is determined according to the planting area in the region. Based on the simulated yields, we calculated the net insurance losses and estimated the systemic risk inherent in insurance portfolios at different aggregated levels.

5.1 Diversification effect across the crops within each country

We first investigated the risk-reducing effect of diversifying among the five crops (corn, cotton, rice, soybeans, and wheat) in the US. We assessed the diversification effect by comparing the net insurance loss under three different scenarios: insuring the five crops separately, insuring them jointly under the estimated correlation structure, and insuring them jointly under the assumption that the indemnities were independent of each other for the five crops (benchmark). Table 1 presents the statistics of the distribution of net insurance income (the opposite of net insurance loss) under the three scenarios. Recall that we used two kinds of clustering copulas (survival Gumbel and Clayton) to model the within-country correlations. We report the simulation results obtained from both copula models.

Since we are assuming fair premiums and ignoring other transaction charges, the mean of the net insurance income is zero for all scenarios. At 70% coverage level, the standard deviation is decreased from \$2.45/acre to \$1.77/acre by insuring the five crops jointly, while it would be decreased to \$1.54/acre if assuming independence, as implied by a survival Gumbel-clustering copula model. Under the same copula model, the minimum observed net income, and value at risk at 1% and 5% levels ($\text{VaR}(1\%)$ and $\text{VaR}(5\%)$, respectively) have all been improved by combining the insurance policies across the five crops, but to a much smaller extent compared to the independence scenario unless the net losses are very low. At 90% coverage level, the results are consistent with those at 70% coverage level. These results reveal that there is some diversification effect across multiple crops, but it is less significant than if risks were independent among different crops. This indicates that there exist some positive correlations of yields among these crops.

Rather similar results are obtained under a Clayton-clustering copula model, showing that the statistics are robust. The only big difference of the results between these two copula

models is that the Clayton model implies a much greater maximum net loss than the survival Gumbel model for all scenarios under both coverage levels. However, the diversification effect (DE), which is defined as the percentage of the risk diversified off with the joint insurance portfolio, is not that different between the two copula models even for the maximum net loss, as reported in Table 2. We calculated the DE of insuring the crops jointly with the estimated correlations, as well as the DE of insuring jointly assuming independence for comparison. By taking a close look at the results, we found a pattern of the gap between the DE under the estimated correlation structure and the DE under independence assumption. For relatively small risks, such as losses around \$2/acre, the gap hardly exists. As the risks become larger, the gap is more and more obvious. For example, for losses around \$10/acre, the gap is about 10%. For losses around \$50/acre, the gap becomes about 20%. When losses are more than \$100/acre, the gap almost reaches 50%. This indicates that while diversifying relatively small risks across multiple crops is effective, it becomes more and more ineffective when risks are expanding. This result also proves the systemic nature of the risks inherent in crop yields within a certain area, even for different kinds of crops. In summary, diversifying the systemic risk across multiple crops within the US works well for relatively small risks. Relatively large risks can only be diversified partly. The larger the risks, the smaller part of them can be diversified.

Table 1: **Statistics of simulated distribution of net insurance income over five crops in the US, expressed in \$/acre**

	Survival Gumbel				Clayton			
	Std	Min	VaR(1%)	VaR(5%)	Std	Min	VaR(1%)	VaR(5%)
CL 70%								
Separately	2.45	-59.12	-10.40	-2.05	2.95	-103.64	-10.74	-1.87
Jointly	1.77	-52.66	-7.01	-1.68	2.41	-98.46	-7.42	-1.71
Jointly (ind.)	1.54	-38.59	-6.09	-1.68	1.76	-50.75	-6.52	-1.72
CL 90%								
Separately	12.47	-134.48	-52.72	-24.67	12.97	-180.92	-53.48	-25.32
Jointly	9.64	-122.73	-41.15	-18.05	10.44	-175.69	-42.40	-19.38
Jointly (ind.)	8.01	-77.40	-32.54	-15.53	8.02	-90.28	-32.08	-15.45

Table 2: **Diversification effect (DE) by combining insurance contracts across five crops in the US**

	Survival Gumbel				Clayton			
	Std	Min	VaR(1%)	VaR(5%)	Std	Min	VaR(1%)	VaR(5%)
CL 70%								
DE	0.277	0.109	0.326	0.183	0.183	0.050	0.309	0.082
DE (ind.)	0.374	0.347	0.414	0.182	0.406	0.510	0.393	0.080
CL 90%								
DE	0.227	0.087	0.219	0.268	0.195	0.029	0.207	0.235
DE (ind.)	0.357	0.424	0.383	0.370	0.381	0.501	0.400	0.390

Similar results were found for the risk-reducing effect by diversifying across the five crops in China (see Table 3 and Table 4). The risks have been decreased by combining insurance policies. However, just the same as in the US, the diversification effect is not as significant as if risks were independent among the crops in China. This again demonstrates the positive correlations of crop yields resulted from correlated weather pattern within a country. The same pattern of diversification effect has also been found in the case of China, that is,

risks are more difficult to diversify across multiple crops if the risks are in relatively larger scales. Thus, it would be of interest to investigate whether large risks can be diversified more effectively across different countries.

Table 3: **Statistics of simulated distribution of net insurance income over five crops in China, expressed in \$/acre**

	Survival Gumbel				Clayton			
	Std	Min	VaR(1%)	VaR(5%)	Std	Main	VaR(1%)	VaR(5%)
CL 70%								
Separately	1.75	-35.89	-8.07	-1.92	2.54	-96.40	-8.83	-1.85
Jointly	1.17	-31.06	-4.94	-1.72	1.98	-80.73	-5.32	-1.75
Jointly (ind.)	0.89	-13.49	-3.78	-1.70	1.24	-30.36	-4.66	-1.81
CL 90%								
Separately	8.89	-94.90	-36.25	-17.80	9.87	-180.13	-38.66	-19.02
Jointly	6.79	-86.90	-29.17	-12.24	7.66	-163.15	-29.10	-12.82
Jointly (ind.)	4.21	-34.73	-14.43	-8.14	4.68	-52.27	-15.41	-8.72

Table 4: **Diversification effect (DE) of combining insurance contracts across five crops in China**

	Survival Gumbel				Clayton			
	Std	Max	VaR(1%)	VaR(5%)	Std	Max	VaR(1%)	VaR(5%)
CL 70%								
DE	0.329	0.134	0.388	0.103	0.222	0.163	0.397	0.057
DE (ind.)	0.491	0.624	0.532	0.116	0.513	0.685	0.473	0.021
CL 90%								
DE	0.237	0.084	0.195	0.312	0.224	0.094	0.247	0.326
DE (ind.)	0.527	0.634	0.602	0.542	0.526	0.710	0.601	0.541

5.2 Diversification effect across the countries

To investigate the risk-reducing effect of diversifying across the US and China, we still compare the net insurance loss under different scenarios: insuring the crops in the US and China separately, insuring the crops in these two countries jointly under the estimated correlation structure, and insuring them jointly under the assumption that the indemnities were independent between the US and China (benchmark). Table 5 and Table 6 shows the simulation results for the distribution of net insurance income and the diversification effect (DE) under different scenarios, respectively. Recall that we also used two kinds of nesting copulas (Gaussian and Frank) to model the between-country correlation. Along with the two kinds of clustering copulas, there are 4 (2×2) different copula models for the whole correlation structure of all the crop yields in the two countries. We report the results obtained from all the four copula models.

The risks have been significantly reduced by diversifying across the two countries. For example, at 90% coverage level and with Clayton clustering copulas, the standard deviation is decreased from \$9.00/acre to \$6.20/acre with Gaussian nesting copula, and to \$6.10/acre with Frank nesting copula. The minimum observed net income is increased from a loss of \$169.18/acre to a loss of \$100.01/acre for Gaussian nesting, and to a loss of \$81.86/acre for Frank nesting. Value at risk at 1% (5%) of net income is improved from a loss of \$35.50/acre (\$15.97/acre) to a loss of \$23.69/acre (\$11.66/acre) for Gaussian nesting, and to a loss of \$23.43/acre (\$11.65/acre) for Frank nesting. Consistent results are obtained from all other copula models at both coverage levels. The most striking is that the diversification effect across countries is comparable to that if risks were independent between the US and China, in terms of all standard deviation, minimum net income, and value at risk. That is, the

diversification effect is quite significant no matter how large the risks are. Different from diversifying across crops, diversifying across countries seems to remove the systemic nature of the risks associated with crop yields.

Another thing to note is that the results do not change much for different copula models. The only big difference, as indicated before, is that the Clayton clusering model implies a larger maximum net loss than the survival Gumbel clusering model. However, the difference has disappeared when calculating the diversification effect. This robustness suggests that the results would not be affected much with different kinds copula modeling, which confirms the stability and reliability of the estimated correlations among crop yields.

Table 5: **Statistics of simulated distribution of net insurance income over the US and China, expressed in \$/acre**

	Survival Gumbel clustering				Clayton clustering			
	Std	Min	VaR(1%)	VaR(5%)	Std	Min	VaR(1%)	VaR(5%)
CL								
70%								
Separately	1.46	-37.40	-5.93	-1.70	2.19	-89.25	-6.33	-1.73
Jointly								
(Gaussian	1.09	-23.50	-4.61	-1.47	1.44	-51.44	-4.89	-1.53
nesting)								
Jointly								
(Frank	1.13	-25.32	-4.63	-1.44	1.35	-47.25	-4.83	-1.48
nesting)								
Jointly								
(ind.)	1.04	-24.47	-4.22	-1.43	1.55	-47.09	-4.80	-1.51
CL								
90%								
Separately	8.16	-104.13	-34.93	-15.03	9.00	-169.18	-35.50	-15.97
Jointly								
(Gaussian	6.02	-64.30	-23.39	-11.89	6.20	-100.01	-23.69	-11.66
nesting)								
Jointly								
(Frank	6.05	-60.83	-24.01	-11.68	6.10	-81.86	-23.43	-11.65
nesting)								
Jointly								
(ind.)	5.82	-61.12	-22.38	-11.28	6.42	-110.10	-24.08	-11.81

Table 6: **Diversification effect (DE) of combining insurance contracts across the US and China**

	Survival Gumbel clustering				Clayton clustering			
	Std	Min	VaR(1%)	VaR(5%)	Std	Min	VaR(1%)	VaR(5%)
CL								
70%								
DE								
(Gaussian nesting)	0.281	0.405	0.259	0.159	0.287	0.410	0.247	0.108
DE								
(Frank nesting)	0.279	0.389	0.254	0.180	0.292	0.406	0.218	0.138
DE								
(ind.)	0.285	0.346	0.289	0.161	0.289	0.472	0.242	0.126
CL								
90%								
DE								
(Gaussian nesting)	0.288	0.343	0.355	0.225	0.293	0.401	0.329	0.268
DE								
(Frank nesting)	0.283	0.429	0.329	0.237	0.298	0.475	0.348	0.256
DE								
(ind.)	0.287	0.413	0.359	0.249	0.287	0.349	0.322	0.261

6 A Sensitivity Analysis for the Selection of the Nesting Copula

The simulation results indicate that diversifying the yield risks across countries has removed the systemic nature of the risks. Rather than recognizing this as suggested by the historical

yield data, it may be argued that it is because the selected nesting copulas, which represent the between-country correlation, imply no lower-tail dependence. Because of the long geographic distance between the U.S. and China as well as the plots of the historical yields (Figure 2), it seems reasonable to assume the weather patterns are not correlated in these two countries and there is no tail-dependence for the correlation of the yields between the two countries. However, to rule out the possibility that the selection of the nesting copula has an effect on the diversification effect, we model the between-country correlation by choosing nesting copulas that imply lower-tail dependence in this section. The simulation results are compared to the results obtained before using nesting copulas that imply no tail dependence.

Results from using three alternative types of nesting copula with survival Gumbel clustering are reported in Table 7. The Gaussian and Frank nesting copulas are the ones used before as they imply no tail dependence. The survival Gumbel nesting copula is the one that implies lower-tail dependence. From Table 7, the summary statistics and diversification effect from the survival Gumbel nesting copula are almost the same as those from the other two nesting copulas, which are also comparable to the independent case. Similar results are found for the Clayton clustering case. As reported in Table 8, the Clayton nesting, which implies lower-tail dependence, results in no significant differences in summary statistics and diversification effect from the other two nesting types and the independent case.

The sensitivity analysis suggests that results obtained before are quite resilient to alternative nesting copulas. Even if a nesting copula that allows for lower-tail dependence is selected, alike diversification effect has been found and no systemic nature has been observed in the yield risks between the two countries. This confirms the assumption that no tail dependence exists for the between-country correlation and makes it convincing for the significant diversification effect across the two countries.

Table 7: **Comparison of the simulation results from different types of nesting copula with survival Gumbel clustering**

	Statistics of net insurance income				Diversification effect			
	Std	Min	VaR(1%)	VaR(5%)	Std	Min	VaR(1%)	VaR(5%)
CL								
70%								
Survival								
Gumbel	1.13	-22.47	-4.77	-1.47	0.282	0.387	0.232	0.158
nesting								
Gaussian	1.09	-23.50	-4.61	-1.47	0.281	0.405	0.259	0.159
nesting								
Frank	1.13	-25.32	-4.63	-1.44	0.279	0.389	0.254	0.180
nesting								
Ind.	1.04	-24.47	-4.22	-1.43	0.285	0.346	0.289	0.161
CL								
90%								
Survival								
Gumbel	6.01	-60.49	-23.67	-11.79	0.285	0.343	0.324	0.235
nesting								
Gaussian	6.02	-64.30	-23.39	-11.89	0.288	0.343	0.355	0.225
nesting								
Frank	6.05	-60.83	-24.01	-11.68	0.283	0.429	0.329	0.237
nesting								
Ind.	5.82	-61.12	-22.38	-11.28	0.287	0.413	0.359	0.249

Table 8: **Comparison of the simulation results from different nesting copulas with Clayton clustering**

	Statistics of net insurance income				Diversification effect			
	Std	Min	VaR(1%)	VaR(5%)	Std	Min	VaR(1%)	VaR(5%)
CL 70%								
Clayton nesting	1.37	-46.23	-4.68	-1.48	0.289	0.475	0.241	0.143
Gaussian nesting	1.44	-51.44	-4.89	-1.53	0.287	0.410	0.247	0.108
Frank nesting	1.35	-47.25	-4.83	-1.48	0.292	0.406	0.218	0.138
Ind.	1.55	-47.09	-4.80	-1.51	0.289	0.472	0.242	0.126
CL 90%								
Clayton nesting	6.22	-91.29	-23.55	-11.71	0.283	0.454	0.333	0.249
Gaussian nesting	6.20	-100.01	-23.69	-11.66	0.293	0.401	0.329	0.268
Frank nesting	6.10	-81.86	-23.43	-11.65	0.298	0.475	0.348	0.256
Ind.	6.42	-110.10	-24.08	-11.81	0.287	0.349	0.322	0.261

7 Conclusions and Discussions

Systemic risk in agriculture has been inhibiting the establishment of independent private crop insurance markets. Private insurers have to rely on subsidies or maintain prohibitively high reserves for the large portfolio insurance risks. This has been hindering the development of farm sector and lowering the welfare of farmers. This study has provided a potential solution for this problem. The study takes a close look at the effectiveness of diversifying systemic risk

across five crops and two countries. The results have shown significant diversification effect. In addition, diversifying across countries with a long distance seems to remove the systemic nature of the risks. This indicates that the systemic risk can be diversifiable if the risk pool is large enough. Thus, by including more crops and countries, it is possible to eliminate systemic risk because of law of large numbers. One problem could be the availability or accuracy of historical yield data in some other countries. Gathering yield data and eliciting as much reliable information as possible from the possibly scanty data would be challenges to solve in the future.

We apply a modern copula model, the hierarchical Kendall copula (HKC) model, to estimate the correlations among the yield variables. The HKC is superior in that it achieves both flexibility and parsimony when modeling correlations. The flexibility makes the HKC model competent to represent various correlation structures. This can be useful to model the complicated correlation structure among yield variables. However, a question may be raised regarding how we choose the correlation structure that fits the best for the specific problem in hand. In this preliminary study, a few alternative HKC models with different building blocks are tried. They do not lead to much different results at this point. But we still need some model selection criteria to figure out the most suitable one, especially if big differences are implied from alternative models when we include more variables. The parsimony of the HKC makes it quite efficient in modeling the correlations among high-dimensional variables. However, as we are including more and more variables in the future, the curse of dimensionality could be another problem that we have to deal with.

The copula-based correlation modeling methods proposed in this study could also be applied in other financial areas to develop risk management tools and insurance products for both agricultural and non-agricultural applications. It is easy to use the methods to

deal with the data in other financial areas because the copula model separates the marginal distributions and the correlation of joint variables. Thus it can be applied to variables with arbitrary marginal distributions. The flexibility and parsimony of the copula model also ensures that it can represent different kinds of correlation structures among potentially high dimensional variables. Furthermore, estimation risk may play an important role for conducting accurate and efficient risk management. It can be taken into consideration with the proposed Bayesian approach to estimating copula functions in this study.

References

- Babcock, B. A., and D. A. Hennessy (1996). Input Demand under Yield and Revenue Insurance. *American Journal of Agricultural Economics* 78: 416-27.
- Bokusheva, R. (2011). Measuring dependence in joint distributions of yield and weather variables. *Agricultural Finance Review*, 71:120-141.
- Brechmann, E. C. (2013). Hierarchical Kendall copulas: Properties and inference. *Canadian Journal of Statistics*.
- Brooks, S.P. (1998). Markov chain Monte Carlo method and its application. *The Statistician*, Vol. 47, pp. 69-100.
- Devroye, L. (1986). Sample-based non-uniform random variate generation. In *Proceedings of the 18th conference on Winter simulation* (pp. 260-265). ACM.
- Embrechts, P., McNeil, A. J. and Straumann, D. (1999). Correlation and dependence in risk management: Properties and pitfalls. *RISK* pp. 69-71.
- Gelman, A., & Rubin, D. B. (1992). Inference from iterative simulation using multiple sequences. *Statistical science*, 457-472.
- Genest, C., & Rivest, L. P. (1993). Statistical inference procedures for bivariate Archimedean copulas. *Journal of the American statistical Association*, 88(423), 1034-1043.
- Goodwin, B. (2001). Problems With Market Insurance in Agriculture. *American Journal of Agricultural Economics*, 83: 643-649.
- Goodwin, B. K. (2012). Copula-Based Models of Systemic Risk in US Agriculture: Implications for Crop Insurance and Reinsurance Contracts.
- Holly Wang, H., and Zhang, H. (2003). On the possibility of a private crop insurance market: A spatial statistics approach. *Journal of Risk and Insurance*, 70(1), 111-124.

Imlahi, L., Ezzer, M., & Chakak, A. (1999). Estimación de la curva mediana de una cópula $C(x_1, \dots, x_m)$. *Revista de la Real Academia de Ciencias Exactas, Físicas y Naturales*, 93(2), 241-250.

Joe, H. (2005). Asymptotic efficiency of the two-stage estimation method for copula-based models. *Journal of Multivariate Analysis*, 94(2), 401-419.

Joe, H., & Xu, J. J. (1996). The estimation method of inference functions for margins for multivariate models (Vol. 166). Technical report.

Junker, M. and May, A. (2005). Measurement of aggregate risk with copulas. *Econometrics Journal* 8: 428-454.

McNeil, A., R. Frey, and P. Embrechts (2005). *Quantitative Risk Management*. Princeton, NJ: Princeton University Press.

Miranda, M. J. and J. Glauber (1997). Systemic Risk, Reinsurance, and the Failure of Crop Insurance Markets. *American Journal of Agricultural Economics*: 206-15.

Nelsen, R. B. (2006). *An Introduction to Copulas*. New York: Springer-Verlag.

Nelson, C. H., and P. V. Preckel (1989). The Conditional Beta Distribution as a Stochastic Production Function. *American Journal of Agricultural Economics* 71: 370S78.

Okhrin, O. (2010). Fitting High-Dimensional Copulae to Data. SFB 649 Discussion Paper, 2010-022, Humboldt-Universität zu Berlin, Germany.

Okhrin, O., Odening, M. and Xu, W. (2012). Systemic weather risk and crop insurance: the case of China. *Journal of Risk and Insurance*, Vol. 80 No. 2, pp. 351-372.

Okhrin, O., Ristig, A. (2012). Hierarchical Archimedean Copulae: The HAC Package. SFB 649 Discussion Papers SFB649DP2012-036. Sonderforschungs- bereich 649, Humboldt University, Berlin, Germany.

Ramirez, O. A. (1997). Estimation and Use of a Multivariate Parametric Model for

Simulating Heteroskedastic, Correlated, Nonnormal Random Variables: The Case of Corn Belt Corn, Soybeans, and Wheat Yields. *American Journal of Agricultural Economics* 79: 191S205.

Smith, M. S. (2011). Bayesian approaches to copula modelling.

Sklar, A. (1959). Fonctions de re´partition a´ n dimension et leurs marges. *Publications de l’Institut de Statistique de l’Universite´ de Paris*, Vol. 8, pp. 229-31.

Trivedi, P.K. and Zimmer, D.M. (2005). Copula modeling: an introduction to practitioners. *Foundations and Trends in Econometrics*, Vol. 1 No. 1, pp. 1-111.

Turvey, C., Nayak, G., & Sparling, D. (1999). Reinsuring agricultural risk. *Canadian Journal of Agricultural Economics/Revue canadienne d’agroeconomie*, 47(3), 281-291.

Vedenov, D. (2008). Application of copulas to estimation of joint crop yield distributions. In *Proceedings of the American Agricultural Economics Association Annual Meeting*, Orlando, FL.

Holly Wang, H., and Hao Zhang (2003). On the possibility of a private crop insurance market: A spatial statistics approach.” *Journal of Risk and Insurance* 70, no. 1: 111-124.

Xu, W., G. Filler, M. Odening, and O. Okhrin (2010). On the Systemic Nature of Weather Risk. *Agricultural Finance Review*, 70: 267-284.

Yan, J. (2007). Enjoy the Joy of Copulas: With a Package copula. *Journal of Statistical Software*, Vol. 21, Issue 4.

Yan, J. and I. Kojadinovic (2012). Package “copula”: Multivariate Dependence With Copulas. R-Project CRAN Repository, February, 2012.

Zhu, Y., S. Ghosh, and B. Goodwin (2008). Modeling Dependence in the Design of Whole Farm Insurance Contract—A Copula-based Approach. Selected paper at the Annual Meeting of the AAEA, 2008.